

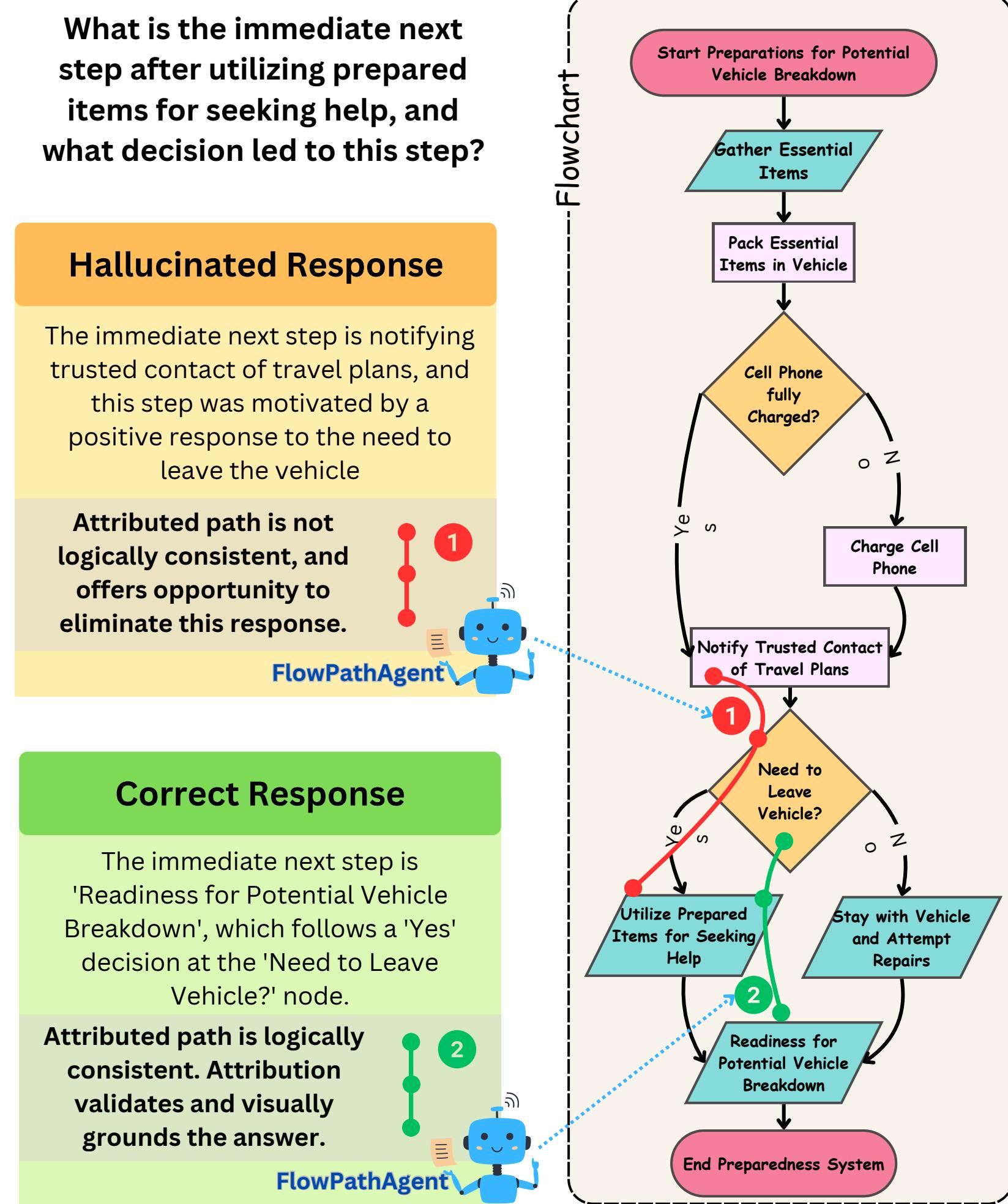
FOLLOW THE FLOW: FINE-GRAINED FLOWCHART ATTRIBUTION WITH NEUROSYMBOLIC AGENTS

Manan Suri, Puneet Mathur, Nedim Lipka,
Franck Dernoncourt, Ryan A Rossi, Vivek Gupta, Dinesh Manocha



ATTRIBUTION IN FLOWCHARTS

Attribution (represented by with FlowPathAgent ensures logical consistency in flowchart-based reasoning. FlowPathAgent uses a neurosymbolic approach to generate attribution paths (&) in the flowchart. This enhances interpretability and reliability in flowchart driven automated decision-making



TASK DEFINITION

Goal:

Given a flowchart image c_i and a natural-language statement s_i , find the sequence of visual regions \mathcal{R}_{s_i} grounding s_i in the image.

Formulation:

$$F : (c_i, s_i) \mapsto \mathcal{R}_{s_i}$$

- Each flowchart $c_i \in \mathcal{F}$ has a graph $G_i = (V_i, E_i)$
- V_i : logical nodes (operations/statements)
- E_i : edges (flow between operations)

Output:

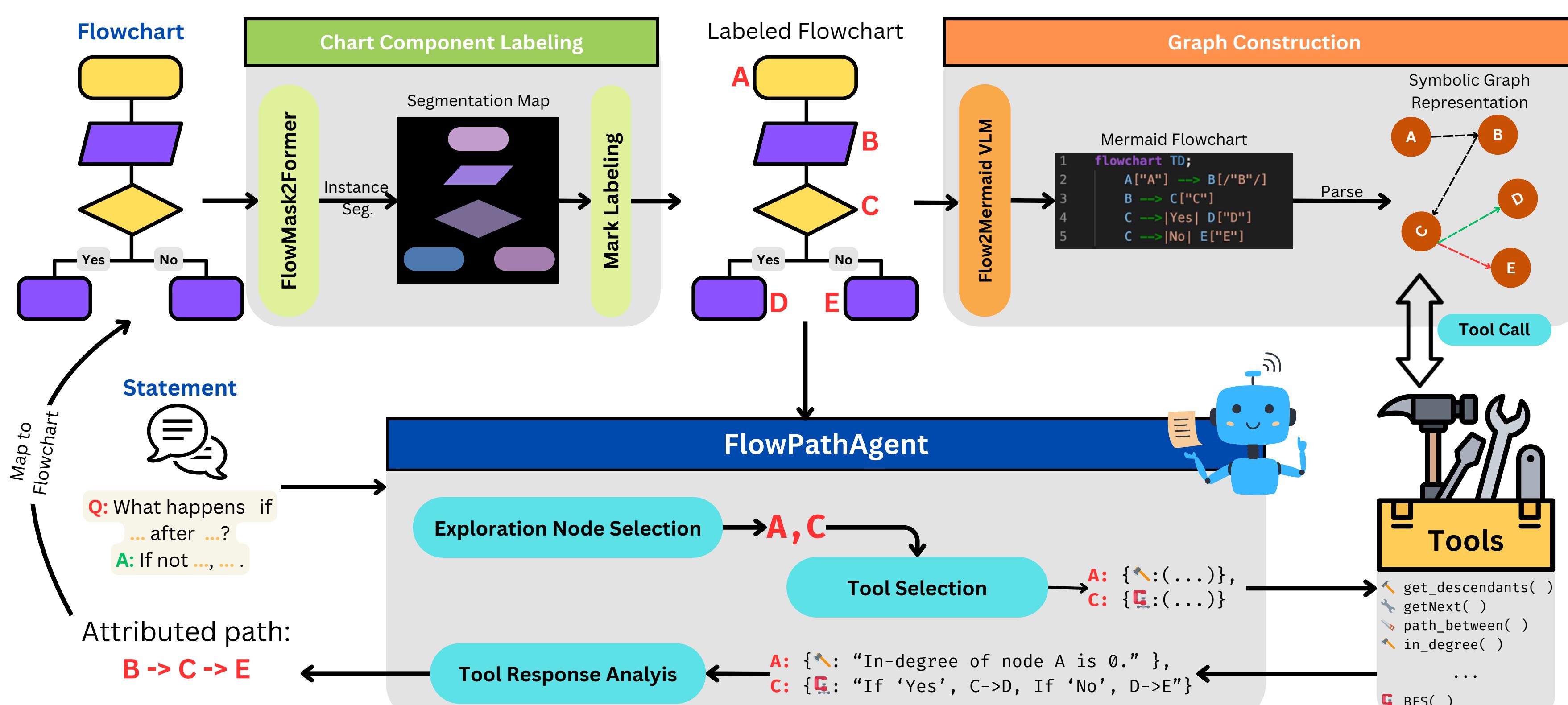
$$\mathcal{R}_{s_i} = \{r_{i1}, r_{i2}, \dots, r_{in}\}$$

Sequence of regions corresponding to a path of logical nodes grounding s_i .

Criteria:

- Optimality:** Shortest path grounding s_i .
- Contextual Alignment:** Matches actions/decisions in s_i .
- Exclusivity:** No extra regions needed beyond \mathcal{R}_{s_i} .

FLOW PATH AGENT

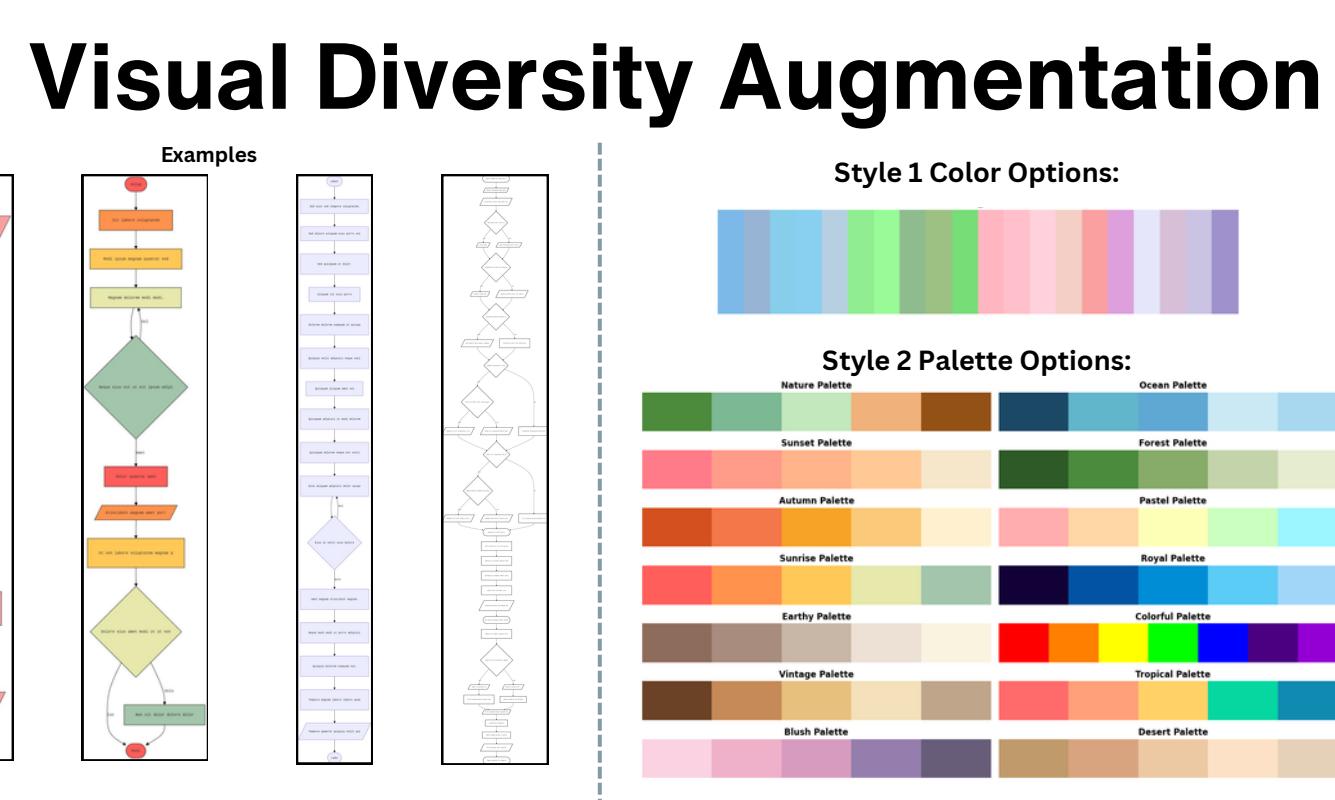


FLOW EXPLAIN BENCH

Data Source

• FlowVQA Dataset

Singh, Shubhankar, et al. "FlowVQA: Mapping multimodal logic in visual question answering with flowcharts." *arXiv preprint arXiv:2406.19237* (2024).



Attribution Annotation

- Automatic Labeling
- Human Verification
- Question Filtering

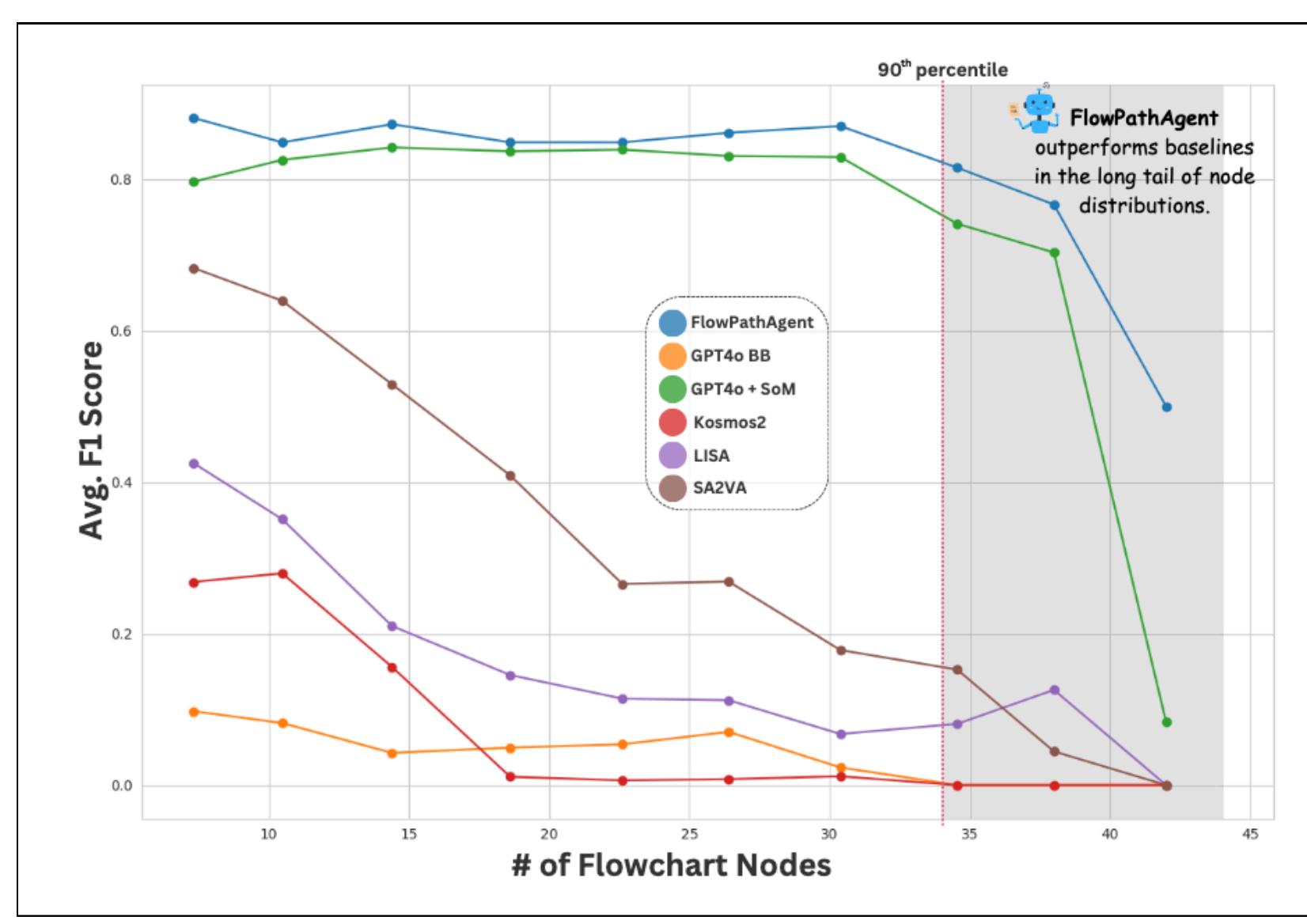
Dataset Composition

	Code	Wiki	Instruct	Overall
# of Flowcharts	189	470	294	953
# of Questions	246	610	382	1238
Fact Retrieval	88	163	102	353
Applied Scenario	69	128	90	287
Flow Referential	43	128	87	258
Topological	46	191	103	340
Avg # of Nodes	11.85	24.49	21.59	21.08
Max # of Nodes	29	43	44	44
Avg Attributed Path Length	2.59	3.21	2.88	2.99
Max Attributed Path Length	15	35	21	35
Avg Words (Question)	26.99	26.12	26.56	26.43
Avg Words (Answer)	8.62	8.74	9.50	8.95

RESULTS

Baseline	Overall			FEBench-Code			FEBench-Wiki			FEBench-Instruct		
	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1
Kosmos-2 (Peng et al., 2023)	37.14	1.76	3.36	41.41	6.45	11.16	20.69	0.31	0.60	38.30	1.64	3.14
LISA (Lai et al., 2024)	18.01	14.34	15.97	35.36	19.18	24.87	14.09	11.74	12.81	18.45	16.18	17.24
SA2VA (Yuan et al., 2025)	66.36	9.88	17.20	79.35	19.34	31.10	58.47	7.40	13.14	65.99	8.82	15.56
VisProg (Gupta and Kembhavi, 2022)	45.95	0.46	0.91	46.88	2.30	4.49	0.00	0.00	0.00	25.00	0.09	0.18
GPT4o Bounding Box	58.82	1.90	3.68	80.00	1.89	3.69	53.19	1.29	2.51	57.89	3.00	5.70
GPT4o SoM	74.10	67.69	70.75	67.32	70.28	68.77	74.55	65.03	69.47	77.84	70.91	74.22
FlowPathAgent	77.19	77.21	77.20	74.18	80.62	77.27	76.29	74.21	75.23	80.28	80.19	80.23

Table 2: Performance comparison of FlowPathAgent with baselines on FlowExplainBench. **Best** and second-best results have been highlighted.



Performance comparison of FlowPathAgent against baselines demonstrates superior effectiveness across long-tail distribution of node count in flowcharts.